RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/764,260

Source: IFW

Date Processed by STIC: 1-3-05

ENTERED

IFWO

```
 3 <110> APPLICANT: Debe, Derek A.
 4        Danzer, Joseph F.
 5        Xie, Lei
 7 <120> TITLE OF INVENTION: METHOD FOR DETERMINING FUNCTIONAL SITES IN A PROTEIN
 9 <130> FILE REFERENCE: 54318.8009.US01
11 <140> CURRENT APPLICATION NUMBER: US 10/764,260
12 <141> CURRENT FILING DATE: 2004-01-22
14 <150> PRIOR APPLICATION NUMBER: US 60/447,562
15 <151> PRIOR FILING DATE: 2003-02-14
17 <160> NUMBER OF SEQ ID NOS: 29
19 <170> SOFTWARE: PatentIn version 3.3
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 327
23 <212> TYPE: PRT
24 <213> ORGANISM: Escherichia coli
26 <400> SEQUENCE: 1
28 Ala Tyr Ile Ala Lys Gln Arg Gln Ile Ser Phe Val Lys Ser His Phe
29 1               5                   10                  15
32 Ser Arg Gln Leu Glu Glu Arg Leu Gly Leu Ile Glu Val Gln Ala Pro
33              20                  25                  30
36 Ile Leu Ser Arg Val Gly Asp Gly Thr Gln Asp Asn Leu Ser Gly Ala
37          35                  40                  45
40 Glu Lys Ala Val Gln Val Lys Val Lys Ala Leu Pro Asp Ala Gln Phe
41      50                  55                  60
44 Glu Val Val His Ser Leu Ala Lys Trp Lys Arg Gln Thr Leu Gly Gln
45 65                  70                  75                  80
48 His Asp Phe Ser Ala Gly Glu Gly Leu Tyr Thr His Met Lys Ala Leu
49              85                  90                  95
52 Arg Pro Asp Glu Asp Arg Leu Ser Pro Leu His Ser Val Tyr Val Asp
53              100                 105                 110
56 Gln Trp Asp Trp Glu Arg Val Met Gly Asp Gly Glu Arg Gln Phe Ser
57          115                 120                 125
60 Thr Leu Lys Ser Thr Val Glu Ala Ile Trp Ala Gly Ile Lys Ala Thr
61      130                 135                 140
64 Glu Ala Ala Val Ser Glu Glu Phe Gly Leu Ala Pro Phe Leu Pro Asp
65 145                 150                 155                 160
68 Gln Ile His Phe Val His Ser Gln Glu Leu Leu Ser Arg Tyr Pro Asp
69              165                 170                 175
72 Leu Asp Ala Lys Gly Arg Glu Arg Ala Ile Ala Lys Asp Leu Gly Ala
73          180                 185                 190
76 Val Phe Leu Val Gly Ile Gly Gly Lys Leu Ser Asp Gly His Arg His
77          195                 200                 205
80 Asp Val Arg Ala Pro Asp Tyr Asp Asp Trp Ser Thr Pro Ser Glu Leu
```

RAW SEQUENCE LISTING   DATE: 01/03/2005
PATENT APPLICATION: US/10/764,260   TIME: 15:40:27

Input Set : A:\54318.8009.US01.txt
Output Set: N:\CRF4\01032005\J764260.raw

```
81       210                215                 220
84  Gly His Ala Gly Leu Asn Gly Asp Ile Leu Val Trp Asn Pro Val Leu
85  225                230                235                 240
88  Glu Asp Ala Phe Glu Leu Ser Ser Met Gly Ile Arg Val Asp Ala Asp
89                  245                250                255
92  Thr Leu Lys His Gln Leu Ala Leu Thr Gly Asp Glu Asp Arg Leu Glu
93                  260                265                270
96  Leu Glu Trp His Gln Ala Leu Leu Arg Gly Glu Met Pro Gln Thr Ile
97           275                280                285
100 Gly Gly Gly Ile Gly Gln Ser Arg Leu Thr Met Leu Leu Leu Gln Leu
101      290                295                300
104 Pro His Ile Gly Gln Val Gln Ala Gly Val Trp Pro Ala Ala Val Arg
105 305                310                315                 320
108 Glu Ser Val Pro Ser Leu Leu
109                 325
112 <210> SEQ ID NO: 2
113 <211> LENGTH: 327
114 <212> TYPE: PRT
115 <213> ORGANISM: Shigella flexneri
117 <400> SEQUENCE: 2
119 Ala Tyr Ile Ala Lys Gln Arg Gln Ile Ser Phe Val Lys Ser His Phe
120 1                5                  10                 15
123 Ser Arg Gln Leu Glu Glu Arg Leu Gly Leu Ile Glu Val Gln Ala Pro
124             20                25                 30
127 Ile Leu Ser Arg Val Gly Asp Gly Thr Gln Asp Asn Leu Ser Gly Cys
128          35                40                 45
131 Glu Lys Ala Val Gln Val Lys Val Lys Ala Leu Pro Asp Ala Gln Phe
132      50                55                 60
135 Glu Val Val His Ser Leu Ala Lys Trp Lys Arg Gln Thr Leu Gly Gln
136 65                70                 75                 80
139 His Asp Phe Ser Ala Gly Glu Gly Leu Tyr Thr His Met Lys Ala Leu
140              85                 90                 95
143 Arg Pro Asp Glu Asp Arg Leu Ser Pro Leu His Ser Val Tyr Val Asp
144             100                105                110
147 Gln Trp Asp Trp Glu Arg Val Met Gly Asp Gly Glu Arg Gln Leu Ser
148          115                120                125
151 Thr Leu Lys Ser Thr Val Glu Ala Ile Trp Ala Gly Ile Lys Ala Thr
152      130                135                140
155 Glu Ala Ala Val Asn Glu Glu Phe Gly Leu Ala Pro Phe Leu Pro Asp
156 145                150                155                160
159 Gln Ile His Phe Val His Ser Gln Glu Leu Leu Ser Arg Tyr Pro Asp
160             165                170                175
163 Leu Asp Ala Lys Gly Arg Glu Arg Ala Ile Ala Lys Asp Leu Gly Ala
164          180                185                190
167 Val Phe Leu Val Gly Ile Gly Gly Lys Leu Ser Asp Gly His Arg His
168          195                200                205
171 Asp Val Arg Ala Pro Asp Tyr Asp Asp Trp Ser Thr Pro Ser Glu Leu
172      210                215                220
175 Gly His Ala Gly Leu Asn Gly Asp Ile Leu Val Trp Asn Pro Val Leu
```

```
176 225                    230                235                240
179 Glu Asp Ala Phe Glu Leu Ser Ser Met Gly Ile Arg Val Asp Ala Asp
180                    245                250                255
183 Thr Leu Lys His Gln Leu Ala Leu Thr Gly Asp Glu Asp Arg Leu Gln
184                    260                265                270
187 Leu Glu Trp His Gln Ala Leu Leu Arg Gly Glu Met Pro Gln Thr Ile
188           275                280                285
191 Gly Gly Gly Ile Gly Gln Ser Arg Leu Thr Met Leu Leu Leu Gln Leu
192       290                295                300
195 Pro His Ile Gly Gln Val Gln Cys Gly Val Trp Pro Ala Ala Val Arg
196 305                    310                315                320
199 Glu Ser Val Pro Ser Leu Leu
200                    325
203 <210> SEQ ID NO: 3
204 <211> LENGTH: 327
205 <212> TYPE: PRT
206 <213> ORGANISM: Salmonella enterica
208 <400> SEQUENCE: 3
210 Ala Tyr Ile Ala Lys Gln Arg Gln Ile Ser Phe Val Lys Ser His Phe
211 1               5                  10                 15
214 Ser Arg Gln Leu Glu Glu Arg Leu Gly Leu Ile Glu Val Gln Ala Pro
215                    20                25                 30
218 Ile Leu Ser Arg Val Gly Asp Gly Thr Gln Asp Asn Leu Ser Gly Cys
219           35                 40                 45
222 Glu Lys Ala Val Gln Val Lys Val Lys Ala Leu Pro Asp Ala Gln Phe
223       50                 55                 60
226 Glu Val Val His Ser Leu Ala Lys Trp Lys Arg Gln Thr Leu Gly Gln
227 65                 70                 75                 80
230 His Asp Phe Ser Ala Gly Glu Gly Leu Tyr Thr His Met Lys Ala Leu
231                    85                 90                 95
234 Arg Pro Asp Glu Asp Arg Leu Ser Pro Leu His Ser Val Tyr Val Asp
235                    100                105                110
238 Gln Trp Asp Trp Glu Arg Val Met Gly Asp Gly Glu Arg Gln Phe Ser
239           115                120                125
242 Thr Leu Lys Ser Thr Val Glu Ala Ile Trp Ala Gly Ile Lys Ala Thr
243       130                135                140
246 Glu Ala Glu Val His Lys Gln Phe Gly Leu Ala Pro Phe Leu Pro Asp
247 145                150                155                160
250 Gln Ile His Phe Val His Ser Gln Glu Leu Leu Ala Arg Phe Pro Asp
251           165                170                175
254 Leu Asp Ala Lys Gly Arg Glu Arg Ala Ile Ala Lys Glu Leu Gly Ala
255           180                185                190
258 Val Phe Leu Val Gly Ile Gly Gly Lys Leu Ser Asp Gly Arg Arg His
259           195                200                205
262 Asp Val Arg Ala Pro Asp Tyr Asp Asp Trp Ser Ser Ala Ser Glu Leu
263       210                215                220
266 Gly Tyr Ala Gly Leu Asn Gly Asp Ile Leu Val Trp Asn Pro Val Leu
267 225                230                235                240
270 Glu Asp Ala Phe Glu Leu Ser Ser Met Gly Ile Arg Val Asp Ala Asp
```

RAW SEQUENCE LISTING   DATE: 01/03/2005
PATENT APPLICATION: US/10/764,260   TIME: 15:40:27

Input Set : A:\54318.8009.US01.txt
Output Set: N:\CRF4\01032005\J764260.raw

```
271                  245              250              255
274 Thr Leu Met Arg Gln Leu Ala Leu Thr Gly Asp Glu Asp Arg Leu Gln
275                  260              265              270
278 Leu Glu Trp His Gln Ala Leu Leu Arg Gly Glu Met Pro Gln Thr Ile
279          275              280              285
282 Gly Gly Gly Ile Gly Gln Ser Arg Leu Thr Met Leu Leu Leu Gln Leu
283      290              295              300
286 Pro His Ile Gly Gln Val Gln Cys Gly Val Trp Pro Ala Gln Val Arg
287 305              310              315              320
290 Glu Ser Ile Pro Ala Ile Leu
291                  325
294 <210> SEQ ID NO: 4
295 <211> LENGTH: 327
296 <212> TYPE: PRT
297 <213> ORGANISM: Salmonella typhimurium
299 <400> SEQUENCE: 4
301 Ala Tyr Ile Ala Lys Gln Arg Gln Ile Ser Phe Val Lys Ser His Phe
302 1                5                10               15
305 Ser Arg Gln Leu Glu Glu Arg Leu Gly Leu Ile Glu Val Gln Ala Pro
306              20               25               30
309 Ile Leu Ser Arg Val Gly Asp Gly Thr Gln Asp Asn Leu Ser Gly Cys
310          35               40               45
313 Glu Lys Ala Val Gln Val Lys Val Lys Ala Leu Pro Asp Ala Gln Phe
314      50               55               60
317 Glu Val Val His Ser Leu Ala Lys Trp Lys Arg Gln Thr Leu Gly Gln
318 65               70               75               80
321 His Asp Phe Ser Ala Gly Glu Gly Leu Tyr Thr His Met Lys Ala Leu
322              85               90               95
325 Arg Pro Asp Glu Asp Arg Leu Ser Pro Leu His Ser Val Tyr Val Asp
326              100              105              110
329 Gln Trp Asp Trp Glu Arg Val Met Gly Asp Gly Glu Arg Gln Phe Ser
330          115              120              125
333 Thr Leu Lys Ser Thr Val Glu Ala Ile Trp Ala Gly Ile Lys Ala Thr
334      130              135              140
337 Glu Ala Glu Val His Lys Gln Phe Gly Leu Ala Pro Phe Leu Pro Glu
338 145              150              155              160
341 Gln Ile Gln Phe Val His Ser Gln Glu Leu Leu Ala Arg Phe Pro Asp
342              165              170              175
345 Leu Asp Ala Lys Gly Arg Glu Arg Ala Ile Ala Lys Glu Leu Gly Ala
346          180              185              190
349 Val Phe Leu Val Gly Ile Gly Gly Lys Leu Ser Asp Gly His Arg His
350          195              200              205
353 Asp Val Arg Ala Pro Asp Tyr Asp Asp Trp Ser Ser Ala Ser Glu Leu
354      210              215              220
357 Gly Tyr Ala Gly Leu Asn Gly Asp Ile Leu Val Trp Asn Pro Val Leu
358 225              230              235              240
361 Glu Asp Ala Phe Glu Leu Ser Ser Met Gly Ile Arg Val Asp Ala Asp
362              245              250              255
365 Thr Leu Met Arg Gln Leu Ala Leu Thr Gly Asp Glu Asp Arg Leu Gln
```

```
366              260              265              270
369 Leu Glu Trp His Gln Ala Leu Leu Arg Gly Glu Met Pro Gln Thr Ile
370         275              280              285
373 Gly Gly Gly Ile Gly Gln Ser Arg Leu Thr Met Leu Leu Leu Gln Leu
374     290              295              300
377 Pro His Ile Gly Gln Val Gln Cys Gly Val Trp Pro Ala Gln Val Arg
378 305              310              315              320
381 Glu Ser Ile Pro Ala Ile Leu
382              325
385 <210> SEQ ID NO: 5
386 <211> LENGTH: 327
387 <212> TYPE: PRT
388 <213> ORGANISM: Yersinia pestis
390 <400> SEQUENCE: 5
392 Gln Phe Ile Gln Lys Gln Gln Gln Ile Ser Phe Val Lys Ser Phe Phe
393 1            5                10               15
396 Ser Arg Gln Leu Glu Gln Gln Leu Gly Leu Ile Glu Val Gln Ala Pro
397             20               25               30
400 Ile Leu Ser Arg Val Gly Asp Gly Thr Gln Asp Asn Leu Ser Gly Ser
401         35               40               45
404 Glu Lys Ala Val Gln Val Lys Val Lys Ser Leu Pro Asp Ser Thr Phe
405     50               55               60
408 Glu Val Val His Ser Leu Ala Lys Trp Lys Arg Lys Thr Leu Gly Arg
409 65               70               75               80
412 Phe Asp Phe Gly Ala Asp Gln Gly Val Tyr Thr His Met Lys Ala Leu
413             85               90               95
416 Arg Pro Asp Glu Asp Arg Leu Ser Ala Ile His Ser Val Tyr Val Asp
417             100              105              110
420 Gln Trp Asp Trp Glu Arg Val Met Gly Asp Gly Glu Arg Asn Leu Ala
421         115              120              125
424 Tyr Leu Lys Ser Thr Val Asn Lys Ile Tyr Ala Ala Ile Lys Glu Thr
425     130              135              140
428 Glu Ala Ala Ile Ser Ala Glu Phe Gly Val Lys Pro Phe Leu Pro Asp
429 145              150              155              160
432 His Ile Gln Phe Ile His Ser Glu Ser Leu Arg Ala Arg Phe Pro Asp
433             165              170              175
436 Leu Asp Ala Lys Gly Arg Glu Arg Ala Ile Ala Lys Glu Leu Gly Ala
437             180              185              190
440 Val Phe Leu Ile Gly Ile Gly Gly Lys Leu Ala Asp Gly Gln Ser His
441         195              200              205
444 Asp Val Arg Ala Pro Asp Tyr Asp Asp Trp Thr Ser Pro Ser Ala Glu
445     210              215              220
448 Gly Phe Ser Gly Leu Asn Gly Asp Ile Ile Val Trp Asn Pro Ile Leu
449 225              230              235              240
452 Glu Asp Ala Phe Glu Ile Ser Ser Met Gly Ile Arg Val Asp Ala Glu
453             245              250              255
456 Ala Leu Lys Arg Gln Leu Ala Leu Thr Gly Asp Glu Asp Arg Leu Glu
457         260              265              270
460 Leu Glu Trp His Gln Ser Leu Leu Arg Gly Glu Met Pro Gln Thr Ile
```

VERIFICATION SUMMARY   DATE: 01/03/2005
PATENT APPLICATION: US/10/764,260   TIME: 15:40:28

Input Set : A:\54318.8009.US01.txt
Output Set: N:\CRF4\01032005\J764260.raw